Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 18-21 August 2005

COMPARING THE PERFORMANCE OF DIFFERENT NEURAL NETWORKS ARCHITECTURES FOR THE PREDICTION OF MINERAL PROSPECTIVITY

CHUN CHE FUNG[1], VANAJA IYER[2], WARICK BROWN[3], KOK WAI WONG[1]

[1] Centre for Enterprise Collaboration in Innovative Systems (CECIS), Murdoch University, Perth, Western Australia
[2]School of Information Technology, Murdoch University, Perth, Western Australia
[3] Centre for Global Metallogeny, University of Western Australia, Perth, Western Australia
E-MAIL: [1,2]{viyer, L.Fung, K.wong}@ Murdoch.edu.au, [3]{wbrown@segs.uwa.edu.au}

Abstract:

In the mining industry, effective use of geographic information systems (GIS) to identify new geographic locations that are favorable for mineral exploration is very important. However, definitive prediction of such location is not an easy task. In this paper, four different neural networks, namely, the Polynomial Neural Network (PNN), General Regression Neural Network (GRNN), Probabilistic Neural Network (PrNN) and Back Propagation Neural Network (BPNN) have been used to classify data corresponding to cells in a map grid into deposit cells and barren cells. These approaches were tested on the GIS mineral exploration data from the Kalgoorlie region of Western Australia. The performance of individual neural networks is compared based on simulation results. The results demonstrate various degrees of success for the networks and suggestions on how to integrate the results are discussed.

Keywords:

Mineral Prospectivity; Polynomial Neural Network,; General Progression Neural Network; Probabilistic Neural Network; Backpropagation Neural Network

1. Introduction

Mineral prospectivity prediction is a problem that involves determining the potential of areas in a regional-scale map to contain mineral deposits for exploration purposes. One of the definitive ways of determining if a location contains mineral deposits is to conduct drilling operations in the area. But this is a trial and error process which is costly and risky. The other method is to use data about known deposits from related areas in some prediction algorithm or system. For this purpose, analysis of data sets contained in the Geographic Information Systems (GIS) of a given area is an important part of the task. The advantage of GIS is its capability of handling large spatial data sets from a variety of sources.

With the increasing software and hardware capacities of computers in the last few decades, geographic information systems that can handle large spatial data sets are readily available. Systems that used to be implemented on mainframe computers are now installed on desk-top computers. GIS is defined [2], as having the functional capability to bring together spatial data from a large variety of sources into a single data base, as a series of geographically-located data layers that are co-registered, that is, overlap correctly at all locations. In the mineral prospectivity problem, a number of approaches have been used to combine geoscience data sets including Boolean algebra [7] and index overlay method [11, 2, 14]. While binary and index overlay methods are simple, they have the disadvantage that the information in the map is limited to two or just a small range of classes and they are not suitable to model complex non-linear relationships. Statistical methods such as the multiple linear regression were the earliest methods used in mineral prospectivity mapping [6, 7, 18]. But the method is based on a range of assumptions [21], in particular, the normality of the variables and that a linear relationship exists between the input and output variables. These assumptions however are commonly violated in geoscience data sets.

In order to handle the large dimensionality and non-linear characteristics of many real-life problems, artificial neural networks (ANN) have been extensively used in many other fields of research. However, they have only recently been introduced in the area of mineral exploration [3, 4, 8, 19]. Until recently, back propagation neural networks (BPNN) have made up the majority of the neural network applications. Brown et al have published [3, 4] the results of using Back propagation neural network (BPNN) for mineral prospectivity. Lately, more studies have been conducted to investigate the performance of other neural networks for the application domain as described in this paper. Probabilistic neural networks are used by Singer and Kouda [20] to classify deposits into

0-7803-9091-1/05/$20.00 ©2005 IEEE

deposit types based on the presence or absence of 58 ore and alteration mineral. Singer and Kouda [19] compared the performance of probabilistic neural network with weights-of-evidence methods for prediction of mineral potential and they have found probabilistic neural network performance to be better.

In this paper, four different neural network types are used to classify a set of map grid cells as deposit cells and barren cells. The results for each network type were evaluated separately and compared. In general, unanimous results were for most of the cells. However,, there are situations where conflicting results are obtained. This paper will discuss possible means to resolve this problem and directions for further research.

2. Individual Neural Networks used in this study

Four different neural network architectures have been used in this study. They are: Polynomial Neural Networks (PNN), Probabilistic neural network (PrNN), General Regression Neural Network (GRNN) and Back propagation neural network (BPNN). Each one of them is used for different reasons. Recently, many papers have been published [1,5,13,15,16] on the use of Polynomial Neural Networks (PNN) for a variety of applications with good results. But none of the reported applications are in the field of mineral exploration. The flexible architecture of PNN has made it one of the choices for this study. Probabilistic neural networks (PrNN) have been reported to be suitable for classification problems. Therefore they have been chosen in this study. Similar to Polynomial Neural Networks, General Regression Neural Network (GRNN) are chosen because of one of their characteristics is that it is not necessary to define the number of hidden layers or number of neurons per layer in advance. While conventional nonlinear regression techniques involve *a priori* specification of the structure of the regression equations to yield a best fit for the data, the GRNN overcomes these restrictions by adjusting the surface dimension in which the regression surface resides without constraining it to a specific form. Back propagation neural network (BPNN) are by far the most commonly used neural network in the research community. Hence, they are used in this study as the basis for comparison.

2.1. Polynomial Neural Network

PNN has a flexible architecture with a topology which is not predetermined but developed through learning. The design is based on Group Method of Data Handling (GMDH) which was invented by Ivankhnenko in the late 1960s [9, 10]. Since then, the technique has been enhanced by many others. GMDH was developed as a means for identifying nonlinear relations between input and output variables. As described in [15], the GMDH generates successive layers with complex links that are individual terms of a polynomial equation.

The individual terms generated in the layers are partial descriptions of data (PDs) being the quadratic regression polynomials with two inputs. The first layer is created by computing regressions of the input variables and choosing the best ones for survival. For example, if the first two variables a and b are taken and combined into a simple set of polynomial terms the terms would be (1, a, b, ab). Next, all possible models made from these terms are checked and the one that best satisfies an evaluation criterion is retained. The second layer is created by computing regressions of the values in the previous layer along with the input variables and retaining the best candidates. More layers are built until the network results cease to improve based on a termination criterion. The selection criterion used in this study penalizes the models that become too complex to prevent overtraining.

2.2. General Regression Neural Network

GRNN is a memory-based supervised feed-forward network based on nonlinear regression theory for function estimation. GRNN was originally developed as a statistical method referred to in the literature as Nadaraya-Watson kernel regression. The method was reinvented in 1990 as a GRNN by Donald Specht [17].

GRNN is a 3-layer network that has an input layer, a hidden layer consisting of at least one node for each training pattern, and an output layer. The transfer function for this network type consists of a parameter called the smoothing factor, rather than the learning rate and momentum used in BPNN. Given a training data set and an independent data set, the transfer function is optimised by the selection of a single smoothing factor for all nodes, which is the common spherical or radial basis function kernel band width. In most applications there is a unique smoothing factor that produces the minimum Mean Square Error (MSE) between the network output and the desired output. This smoothing factor provides the same service in GRNN as the learning rate and momentum in BPNN determining how tightly the data will match the predictions or fit the curve.

The GRNN network used in this study has 10 inputs and one output. The inputs are the normalized data from the 10 GIS layers corresponding to a single cell in a map grid. The GIS data is described in the next section. A Genetic

Algorithm (GA) is used to find the appropriate individual smoothing factors for each input as well as an overall smoothing factor. Training of the network based on the proposed method proceeds in two parts. The first stage trains the network with a set of training data. The second stage uses a calibration process to test a whole range of smoothing factors. This is intended to obtain a combination that works best on the test set with the network created in the first stage. Compared to BPNN, the GRNN has a fast training time. Results are reported in the following section. The smoothing factor obtained in the training phase is used to calculate outputs for the test data set

2.3. Probabilistic Neural Networks

Probabilistic neural networks are known for their ability to train on sparse data sets. This network separates data into a specified number of output categories. This network is a three-layer network consisting of an input layer to distribute the inputs, a hidden layer with a unit corresponding to each training pattern (grouped according to the class they belong to) and an output layer containing one neuron for each category. In the present study, there are two neurons to represent classification between barren and deposit cells.

2.4. Back Propagation Neural Networks

BPNN is the most popular neural networks and is used in numerous applications. The network is trained using the back propagation algorithm . The learning algorithm performs a gradient descent optimization on the weights linking the nodes in each layer. While BPNN has shortcomings such as long training time and the possibility of over training, it is nevertheless simple to use and has shown to be robust and gives good results in most cases. It is therefore taken as the benchmark with which to compare the other networks.

3. GIS Data Set

A GIS database is viewed as a collection of maps of a particular data type such as solid geology for a common geographic coordinate system. Spatial objects in GIS map layers are stored in either vector or raster format, in which objects are represented as points, lines and polygons or as grid cells, respectively. Most GIS support both structures and allow conversion from one structure to another.

In this study, the GIS data set used is the one described in [4]. The data set was used to predict the prospectivity for orogenic gold deposits in an approximately 100 x 100 km area of the Archean Yilgarn Block, near Kalgoorlie, Western Australia. In this study, 10 GIS layers in raster data format are used to create the input feature vectors. The GIS layers correspond to information such as the solid geology, or the distance to the nearest fault as shown in Figure 1. The thematic layers are divided into a grid of square cells of 100m side. Each cell is represented by the cell position and a single attribute value within the two dimensional matrix of cells. The map area thus results in 1,254,000 cells. Out of these, only 120 cells correspond to deposits with a total gold content exceeding 1000 kg. Together with 148 randomly selected barren or non-deposit cells, they are used as training and test data sets in this study.

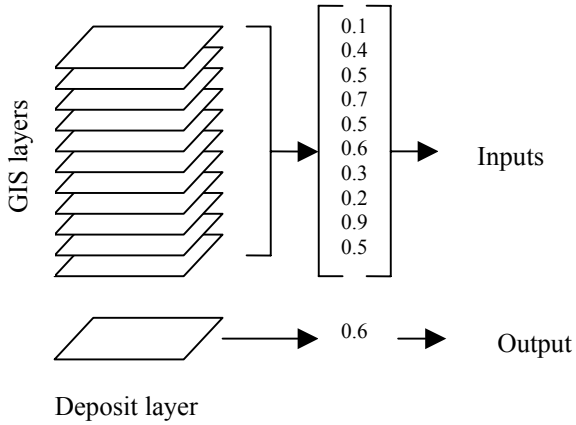

Figure 1: A typical input/output pattern used to train and test the neural networks in this study.

From this total of 268 cells, 187 cells were used for training and 81 cells for testing. Similar to the training data set, the test data set contains both deposit and barren cells. All the input values are scaled to [0, 1]. Table 1 shows the number of patterns in the test and training data sets.

Table 1. Number of patterns in the training and test data sets

Training Data Set			Test Data Set		
Deposit	Barren	Total	Deposit	Barren	Total
85	102	187	35	46	81

For each network architecture, a number of neural networks were trained. The networks comprised of 10-input nodes and a single output node. These networks were initialized with different random weights. Of these, the one which gives the best training set results is tested with the independent test set of 81 patterns.

The output values ranged from 0 to 1. These output

values were classified as barren or deposit cells by comparing the network output against different threshold or cut-off probability values. The output classes were determined using thresholds ranging from 0.1 to 0.9 in steps of 0.1. The results obtained in each case are shown in Tables 2 (a)-(d) below.

Table 2: Comparing results from different networks

Threshold value	% correct Training set		% correct Test set	
	Deposit	**Barren**	**Deposit**	**Barren**
0.1	100.0	**32.4**	100.0	**30.4**
0.2	98.8	53.9	100.0	43.5
0.3	97.6	66.7	91.4	54.3
0.4	90.6	77.5	82.9	65.2
0.5	**81.2**	**87.3**	**80.0**	**71.7**
0.6	71.8	91.2	65.7	78.3
0.7	63.5	95.1	60.0	84.8
0.8	44.7	98.0	37.1	93.5
0.9	**25.9**	99.0	**20.0**	97.8

(a) Polynomial Neural Network (PNN)

Threshold value	% correct Training set		% correct Test set	
	Deposit	**Barren**	**Deposit**	**Barren**
0.1	97.7	**80.4**	88.6	**60.9**
0.2	97.7	88.2	88.6	67.4
0.3	97.7	91.2	88.6	71.7
0.4	96.5	95.1	85.7	71.7
0.5	**96.5**	**96.1**	**82.9**	**73.9**
0.6	91.8	99.0	77.1	78.3
0.7	90.1	99.0	74.3	78.3
0.8	89.4	99.0	62.9	80.4
0.9	**83.5**	99.0	**60.0**	84.8

(b) General Regression Neural Network (GRNN)

Threshold value	% correct Training set		% correct Test set	
	Deposit	**Barren**	**Deposit**	**Barren**
0.1	100	**42.0**	97.1	**37.0**
0.2	100	57.8	88.6	43.5
0.3	98.8	66.7	85.7	52.2
0.4	96.5	81.4	80.0	58.7
0.5	**94.1**	**87.3**	**77.1**	**65.2**
0.6	89.4	91.2	74.3	71.7
0.7	75.3	91.2	68.6	82.6
0.8	61.2	94.1	42.9	89.1
0.9	**47.1**	95.1	**37.1**	93.5

(c) Results from Backpropagation Neural Network (BPNN)

% correct Training set results		% correct Test set results	
Deposit	**Barren**	**Deposit**	**Barren**
89.4	94.1	80.0	71.7

(d) Probabilistic Neural Network (PrNN)

In the above investigation, outputs from the PNN, GRNN and BPNN are within the range of 0 to 1. The threshold value is therefore used to determine whether a cell should be classed as a deposit cell or a barren cell. However, PrNN gives the results in binary form, therefore the threshold value is not included in Table 2(d).

In Tables 2(a)-(c), 0.5 is used as the nominal cut-off. It can be seen that PNN, GRNN and PrNN all outperformed BPNN. If threshold values of 0.9 and 0.1 are used to classify deposit cells and barren cells, respectively, then the GRNN is the best performing network and the PNN is the poorest performing network. The widely differing classification rates are due to fact that different networks assigned different classes to the same cell.

This naturally leads to the question, "given a particular cell with conflicting results from different networks, how should the system decide whether the unknown cell is a deposit cell or a barren cell?" This is similar to a situation in which different decisions or recommendations are offered by multiple experts. The approach to this problem is similar to the use of neural network ensembles. The authors have already developed rule-base system using heuristic knowledge and observations draw from the network outputs, Fuzzy Logic (FL) and a Genetic Algorithm (GA). This work will be described in subsequent reports.

4. Conclusions

In this paper, the problem of the prediction of mineral propsectivity from GIS data sets has been investigated. The performance of four different neural networks has been compared and additional approaches to solve the problem have also been discussed. The four types of neural networks used are: Polynominal Neural Network (PNN), General Regression Neural Network (GRNN), Backpropagation Neural Network (BPNN) and Probabilistic Neural Network (PrNN). PNN, GRNN and PrNN all outperformed BPNN in terms of accuracy and execution time. While there exist situations where conflicting results may be given by the different networks, one approach to solve the problem is to process the results from the multiple networks as an ensemble and apply rules based on heuristic knowledge, fuzzy logic and genetic algorithm.

Acknowledgements

This paper is supported by a grant from the Australian Research Council (ARC) and Vanaja Iyer is a recipient of the Australian Postgraduate Award Industry (APAI) from the same grant.

References

[1] Baker, B., An Inductive Approach to Production-function Modeling: A Comparison of Group Method of Data Handling (GMDH) and Other Neural Network. 1998.

[2] Bonham_Carter, G. F., Geographic Infomation Systems for Scientists: Modelling with GIS. 1994, New York: Pergamon Press.

[3] Brown, W.M., . Gedeon T., Groves D.I., and Barnes, R. G., Artificial neural networks: a new method for mineral prospectivity mapping. Australian Journal of Earth Sciences, 2000 (**47**): p. 757-770.

[4] Brown, W.M., Groves D.I., and Gedeon T., Use of noise to augment training data: a neural network method of mineral potential mapping in regions of limited known deposits. Natural Resources Research, 2003. **12**(2): p. 141-151.

[5] Dolenko, S.A., Orlov, Y.V., and Persiantsev, I.G., Practical implementation and use of Group method of data handling (GMDH): Prospects and Problems. in proceedings of ACEDC. 1996. PEDC, University of Plymouth, UK.

[6] Geoffroy, J.D. and Wignall, T.K., Statistical decision in regional exploration: application of regression and Bayesian classification analysis in the southwest Winconsin zinc area. Economic Geology, 1970. **65**: p. 769-777.

[7] George, H. and Bonham_Carter, G.F., Spatial modelling of geological data for gold exploration Star Lake area. Statistical applications in the earth Sciences, 1990 p. 157-169.

[8] Harris, D. and Pan, G.., Mineral favourability mapping:a comparison of artificial neural networks, logical regression and discriminant analysis. Natural Resources Research, 1999. **8**: p. 93-109.

[9] Ivakhnenko, A.G., Polynomial theory of complex systems. IEEE Transactions on Systems, Man and Cybernetics, 1971. **1**(4): p. 364-378.

[10] Ivakhnenko, A.G., The Review of Problems Solvable by Algorithms of the Group Method of Data Handling (GMDH). Pattern Recognition and Image Analysis, 1995. **5**(4): p. 527-535.

[11] Koch, G.S., Howarth, R.J., Scheunemeyer J.H. and Lowry R.K., Uranium resource assessment through statistical analysis of exploration geochemical and other data Economic Geology, 1981. **76**: p. 1056-1066.

[12] Miyagishi, K., Ohsako, M. and Ichihashi, H., Temperature prediction from regional spectral model by Neurofuzzy GMDH.

[13] Oh, Sung-Kwun and Pedrycz, W. The design of self-organizing Polynomial Neural Networks. Information Sciences, 2002. **141**: p. 237-258.

[14] Reddy, R.K.T. and G.S. Koch, A case study of silver exploration in parts of Idaho and Montana using mathematical modelling techniques. Economic Geology, 1988. **83** p. 1008-1014.

[15] Sarycheva, L., Using GMDH for modeling economical indices of mine opening. Systems Analysis Modelling Simulation, 2003. **43**(10): p. 1341-1349.

[16] Sarycheva, L., Using GMDH in ecological and socio-economical monitoring problems. Systems Analysis Modelling Simulation, 2003. **43**(10): p. 1409-1414.

[17] Specht, D.F., *A General Regression Neural Network.* IEEE transactions on Neural Networks, 1991. **2**(6): p. 568-576.

[18] Sinclair, A.J. and G.J. Woodsworth, Multiple regression as a method of estimating exploration potential in an area near Terrace, B. C. Economic Geology, 1970. **65**: p. 998-1003.

[19] Singer, D.A. and Kouda, R., A comparison of the weights-of-evidence method and probabilistic neural networks. Natural Resources Research, 1999. **8**(4): p. 287-298.

[20] Singer, D.A. and Kouda, R., Classification of mineral deposits into types using mineralogy with a probabilistic neural network. Nonrenewable Resources, 1997. **6**: p. 27-32.

[21] Hair, J. F., Anderson, R. E., Tatham, R. L. and Black, W. C., Multivariate Data Analysis. 1998, Upper Saddle River, Prentice Hall.